Quality/Value Relationship for Imperfect Information in the Umbrella Problem

Richard W. Katz; Allan H. Murphy

The American Statistician, Vol. 41, No. 3 (Aug., 1987), 187-189.

Stable URL:
http://links.jstor.org/sici?sici=0003-1305%28198708%2941%3A3%3C187%3AQRFIII%3E2.0.CO%3B2-2

The American Statistician is currently published by American Statistical Association.

Quality/Value Relationship for Imperfect Information in the Umbrella Problem

RICHARD W. KATZ and ALLAN H. MURPHY*

The so-called umbrella problem (or cost-loss ratio situation), in which an individual must decide whether to take an umbrella in the face of uncertainty concerning whether it will rain today, is sometimes used as a textbook example of decision making under uncertainty. This problem is extended to provide a simple demonstration of the way in which the economic value of imperfect information changes as its quality increases.

KEY WORDS: Cost-loss ratio situation; Decision analysis; Forecasting; Value of information.

1. INTRODUCTION

To demonstrate the principles of decision making under uncertainty, textbooks sometimes rely on the so-called *umbrella problem* (e.g., Lindgren 1968; Winkler 1972). This term refers to the situation in which an individual must decide whether to take an umbrella in the face of uncertainty concerning whether it will rain today. Moreover, another version of this problem, known as the *cost-loss ratio situation*, is frequently treated in the meteorological literature in assessing the economic value of weather forecasts (e.g., Murphy 1977; Murphy, Katz, Winkler, and Hsu 1985). In this version of the umbrella problem, the individual must choose between two possible actions, say a_1 (take umbrella) or a_0 (do not take umbrella). The random variable Θ representing the weather has two possible states, say $\Theta = 1$ (rain) or $\Theta = 0$ (no rain). If the umbrella is taken, then the decision maker incurs a cost C, whereas if the umbrella is not taken and it does rain, a loss L is incurred. We take $0 < C < L$ in order to treat a problem with a nontrivial solution.

The decision maker's task is to select the action that minimizes the expected expense (equivalent to maximizing expected return and to maximizing expected utility for a risk-neutral decision maker). This expectation is taken with respect to the information that is available about the likelihood of rain. Such information could range from simply a single prior probability of rain, based on historical observations and commonly termed "climatological" information in the meteorological literature, to imperfect weather forecasts specifying a conditional probability of rain that varies from day to day, to perfect information specifying (with probability 1) that rain will or will not occur on a given day. Information is of value only insofar as it changes the actions that the decision maker takes, leading to a reduction in the expected expense relative to the situation in which the information is not available. For a particular form of imperfect information, we show the way in which the economic value of the information changes as its quality increases.

2. INFORMATION

Climatological (or prior) information consists of a single probability of rain, say $\pi = \Pr\{\Theta = 1\}$, that is available to the decision maker each day. Imperfect information about Θ is assumed to consist of the random variable Z, indicating a forecast of rain ($Z = 1$) or of no rain ($Z = 0$). We denote the conditional (or posterior) probabilities of rain by

$$p_1 = \Pr\{\Theta = 1 | Z = 1\}, \qquad p_0 = \Pr\{\Theta = 1 | Z = 0\}. \tag{2.1}$$

Without loss of generality, we assume that these probabilities satisfy the ordering $0 \leq p_0 \leq \pi \leq p_1 \leq 1$. The predictive (or marginal) probability of a forecast of rain, say $p_Z = \Pr\{Z = 1\}$, is related to the climatological and conditional probabilities of rain by

$$p_Z = (\pi - p_0)/(p_1 - p_0) \tag{2.2}$$

for $p_1 > p_0$. This form of imperfect information can be viewed as being derived from a 2×2 contingency table that records past rainfall forecasts and observations. Alternatively, it could be viewed as a special case of probabilistic forecasts in which only two possible probabilities of rain are ever used.

3. VALUE OF INFORMATION

The economic value of imperfect forecasts is measured relative to the situation in which the decision maker has only climatological information available (e.g., Winkler and Murphy 1985). Let E_C and E_F denote the minimal expected expenses associated with climatological information and imperfect forecasts, respectively. Then the value of imperfect forecasts, say V_F, is given by $V_F = E_C - E_F$, that is, the difference in expected expenses depending on whether or not the forecast variable Z is considered. From a general result of Blackwell (1953) concerning the comparative value of experiments (see also Hilton 1981), $V_F \geq 0$. The fact that imperfect information cannot have negative economic value is also related to notions of refinement and sufficiency (e.g., DeGroot and Fienberg 1982, 1986; Vardeman and Meeden 1983). The economic value of information can, of course, be interpreted as the maximal amount that a decision maker is willing to pay for its receipt.

*Richard W. Katz is Scientist, Environmental and Societal Impacts Group, National Center for Atmospheric Research, Boulder, Colorado 80307. Allan H. Murphy is Professor, Department of Atmospheric Sciences, Oregon State University, Corvallis, Oregon 97331. This work was supported by National Science Foundation Grant ATM-8507495. The National Center for Atmospheric Research is sponsored by the National Science Foundation. The authors thank an associate editor and two referees for comments that helped improve the exposition.

For the case of climatological information alone, a comparison of expected expenses (πL if action a_0 is taken and C if action a_1 is taken) implies that an umbrella should be taken if $\pi > C/L$ (hence the name cost-loss ratio situation). In particular, the minimal expected expense is

$$\begin{aligned} E_C &= \pi L \quad \text{if } \pi \leq C/L \\ &= C \quad \text{if } \pi \geq C/L. \end{aligned} \tag{3.1}$$

For the case of imperfect forecasts, the determination of an optimal policy and the minimal expected expense is accomplished in an analogous manner. It is evident that taking an umbrella is optimal if $\Pr\{\Theta = 1|Z\} > C/L$. Clearly, if $p_0 > C/L$ or if $p_1 < C/L$, the optimal policy using the forecasts is identical to that for climatological information; thus the minimal expected expenses are the same (i.e., $V_F = 0$). Only when $p_0 < C/L < p_1$ will there be a difference in optimal policies. Under this condition, an umbrella should only be taken when rain is forecast (i.e., $Z = 1$) and

$$E_F = (1 - p_Z)p_0 L + p_Z C. \tag{3.2}$$

Taking the difference in expected expenses [(3.1) minus (3.2)],

$$\begin{aligned} V_F &= (1 - p_Z)(C - p_0 L) & \text{if } p_0 \leq C/L \leq \pi \\ &= (\pi - p_0)L - p_Z(C - p_0 L) & \text{if } \pi \leq C/L \leq p_1. \end{aligned} \tag{3.3}$$

4. QUALITY OF INFORMATION

The expression (3.3) for the value of imperfect forecasts depends on two parameters, p_0 and p_1, related to the characteristics of the forecasts [recall that the parameter p_Z is redundant by (2.2)]. To simplify matters further, the relative frequency with which rain is forecast might be constrained to equal the relative frequency of rain; that is, consider the case in which

$$p_Z = \pi. \tag{4.1}$$

This constraint implies that

$$p_0 = (1 - p_1)\pi/(1 - \pi). \tag{4.2}$$

So the parameter p_1 ($\pi \leq p_1 \leq 1$) completely determines the characteristics of the imperfect forecasts. In particular, climatological and perfect information are the special, limiting cases of imperfect forecasts in which $p_1 = \pi$ and $p_1 = 1$, respectively.

It is sometimes convenient to consider a linear transformation of p_1; namely,

$$q = (p_1 - \pi)/(1 - \pi). \tag{4.3}$$

Note that $0 \leq q \leq 1$, with $q = 0$ for climatological information and $q = 1$ for perfect information. Hence we regard q as a relative measure of the "quality" of imperfect forecasts. In particular, it is straightforward to show that q is the ordinary correlation coefficient between Θ and Z. The expected Brier score, as well as its decomposition into calibration and refinement components (Blattenberger and Lad 1985; DeGroot and Fienberg 1982, 1986), can also be expressed in terms of the square of q. It is important to observe that the parameters p_1 and q depend only on the probabilistic characteristics of the weather event and forecasts, not on any economic parameters involved in the consequences of the umbrella decision-making problem (i.e., C and L).

5. QUALITY/VALUE RELATIONSHIP

Using (4.1) and (4.2), the expression (3.3) for the value of imperfect forecasts simplifies to

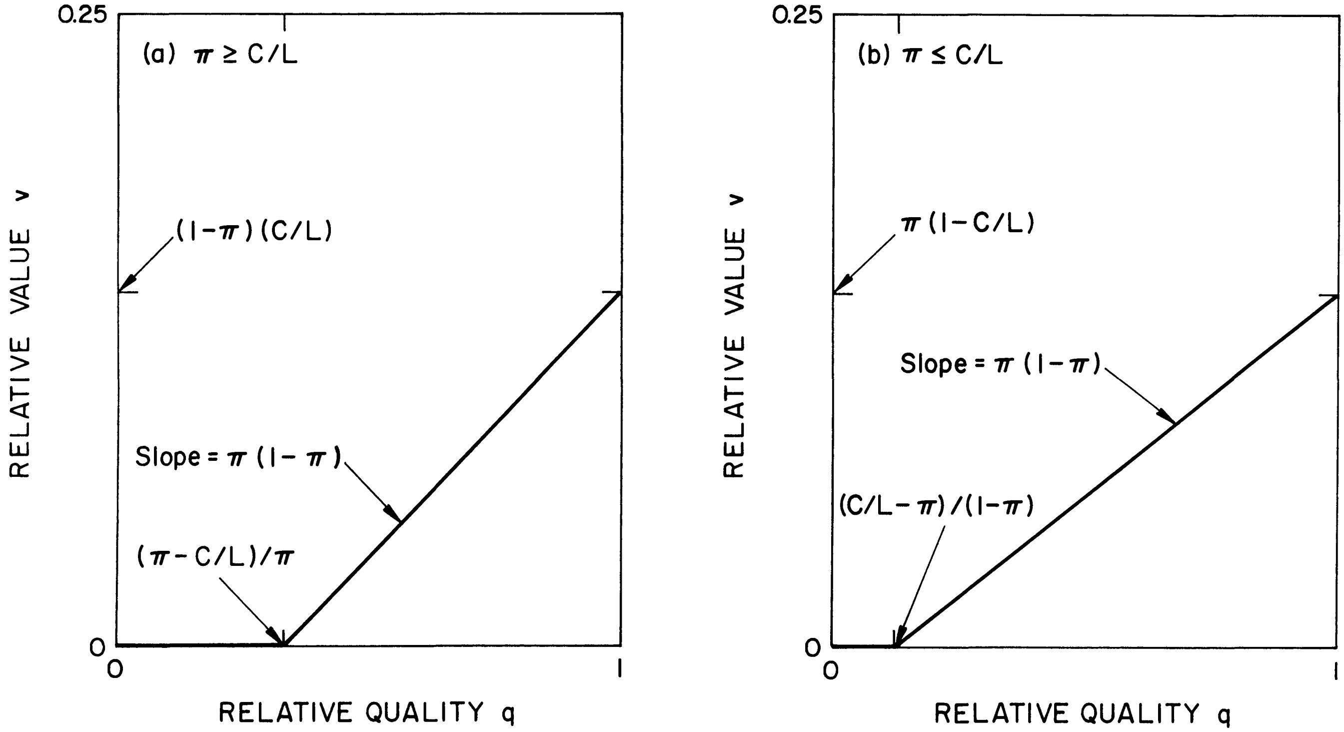

Figure 1. Quality/Value Relationship: Relative Economic Value as a Function of Relative Quality of Imperfect Weather Forecasts.

$$V_F = 0 \quad \text{if } \pi \le p_1 \le p_1^*$$
$$= \pi L(p_1 - p_1^*) \quad \text{if } p_1^* \le p_1 \le 1, \qquad (5.1)$$

where

$$p_1^* = 1 - (C/L)[(1 - \pi)/\pi] \quad \text{if } \pi \ge C/L$$
$$= C/L \quad \text{if } \pi \le C/L. \qquad (5.2)$$

Note that in the $\pi \ge C/L$ case, p_1^* is the value of p_1 for which $p_0 = C/L$. Equations (5.1) and (5.2) express economic value V_F as a simple function of the conditional probability of rain p_1.

Figure 1 shows the corresponding relationship between relative value, $v = V_F/L$ say, and relative quality q. Up to a certain quality threshold, forecasts are of no economic value because the action that is optimal remains the same as when only climatological information is available. Above this threshold, the economic value of the forecasts increases linearly until the upper limit of perfect information is reached. The quality threshold depends on the relative distance between the cost-loss ratio C/L and the climatological probability of rain π, with no threshold only in the special case of $\pi = C/L$. The slope of the linear increase in relative economic value above the quality threshold is simply var(Θ), which is a maximum when $\pi = .5$. The relative economic value of perfect information depends on both π and C/L and can never exceed .25.

6. CONCLUDING REMARKS

A closed-form expression for the relationship between the quality and economic value of imperfect forecasts in the umbrella problem has been presented. This result provides a simple example of the nonlinear nature of the relationship and how, in particular, imperfect information of limited quality may be of no value to a decision maker. Although the umbrella problem is certainly an oversimplified decision-making model, the quality/value relationship obtained is quite suggestive of results that can be derived for more complex models.

For instance, a dynamic version of the cost-loss ratio situation in which it is assumed that the loss L can be incurred at most once could be considered. If the criterion of minimizing the total expected expense over a finite horizon is adopted, then the quality/value relationship still includes a quality threshold below which the value of forecasts is 0. Above this threshold, economic value follows a piecewise linear convex curve, with the number of linear segments increasing as the number of occasions on which decisions are made increases (Murphy et al. 1985). If the criterion of minimizing the total discounted expected expense over an infinite horizon is adopted, then economic value follows a continuous convex curve above a quality threshold (Katz and Murphy 1987).

[*Received June 1986. Revised February 1987.*]

REFERENCES

Blackwell, D. (1953), "Equivalent Comparisons of Experiments," *Annals of Mathematical Statistics*, 24, 265–272.

Blattenberger, G., and Lad, F. (1985), "Separating the Brier Score Into Calibration and Refinement Components: A Graphical Exposition," *The American Statistician*, 39, 26–32.

DeGroot, M. H., and Fienberg, S. E. (1982), "Assessing Probability Assessors: Calibration and Refinement," in *Statistical Decision Theory and Related Topics III* (Vol. 1), eds. S. S. Gupta and J. O. Berger, New York: Academic Press, pp. 291–314.

——— (1986), "Comparing Probability Forecasters: Basic Binary Concepts and Multivariate Extensions," in *Bayesian Inference and Decision Techniques: Essays in Honor of Bruno de Finetti*, eds. P. K. Goel and A. Zellner, Amsterdam: Elsevier, pp. 247–264.

Hilton, R. W. (1981), "The Determinants of Information Value: Synthesizing Some General Results," *Management Science*, 27, 57–64.

Katz, R. W., and Murphy, A. H. (1987), "Quality/Value Relationships for Imperfect Forecasts in the Dynamic Cost-Loss Ratio Decision-Making Model," unpublished manuscript.

Lindgren, B. W. (1968), *Statistical Theory* (2nd ed.), London: Macmillan.

Murphy, A. H. (1977), "The Value of Climatological, Categorical and Probabilistic Forecasts in the Cost-Loss Ratio Situation," *Monthly Weather Review*, 105, 803–816.

Murphy, A. H., Katz, R. W., Winkler, R. L., and Hsu, W.-R. (1985), "Repetitive Decision Making and the Value of Forecasts in the Cost-Loss Ratio Situation: A Dynamic Model," *Monthly Weather Review*, 113, 801–813.

Vardeman, S., and Meeden, G. (1983), "Calibration, Sufficiency, and Domination Considerations for Bayesian Probability Assessors," *Journal of the American Statistical Association*, 78, 808–816.

Winkler, R. L. (1972), *Introduction to Bayesian Inference and Decision*, New York: Holt, Rinehart & Winston.

Winkler, R. L., and Murphy, A. H. (1985), "Decision Analysis," in *Probability, Statistics, and Decision Making in the Atmospheric Sciences*, eds. A. H. Murphy and R. W. Katz, Boulder, CO: Westview Press, pp. 493–524.